Review
The Origins of Social Categorization
Zoe Liberman,1,* Amanda L. Woodward,2 and Katherine D. Kinzler3

Forming conceptually-rich social categories helps people to navigate the complex social world by allowing them to reason about the likely thoughts, beliefs, actions, and interactions of others, as guided by group membership. Nevertheless, social categorization often has nefarious consequences. We suggest that the foundation of the human ability to form useful social categories is in place in infancy: social categories guide the inferences infants make about the shared characteristics and social relationships of other people. We also suggest that the ability to form abstract social categories may be separable from the eventual negative downstream consequences of social categorization, including prejudice, discrimination, and stereotyping. Although a tendency to form inductively-rich social categories appears early in ontogeny, prejudice based on each particular category dimension may not be inevitable.

Social Categorization Profoundly Influences Human Social Life
Despite the salience of individuals in social thinking, a large body of work suggests that the tendency to conceive of people as belonging to social categories is automatic [1–3]. Indeed, the ability to group instances into categories and to use category-based knowledge to generate novel inductive inferences is a powerful aspect of human cognition [1,4]. In particular, the capacity to view category members as sharing important, unchanging, and possibly unobservable similarities allows people to efficiently, and perhaps even spontaneously, learn a property of a category and apply it to novel category members [5–9]. When applied to the social domain, forming conceptually-rich categories has obvious functional value – social categories organize our vast knowledge about human attributes and about the complex relationship networks that comprise human social life [10].

Although category formation has many upsides, much of the research on social categorization focuses on its potential downstream negative consequences. Social categorization differs from other forms of categorization in that people tend to place themselves in a category [11], leading them to be partial to members of their own group (ingroup) relative to those from other groups (outgroup) in terms of social preferences, empathic responding, and resource distribution [12–15]. Beyond sheer partiality and greater liking of members of one’s own group, some of the most invidious effects of social categories result from the biased belief systems that social categorization supports – including stereotypes for, essentialist beliefs about, and even dehumanization of members of some social groups [12,13,16,17].

Although prejudice was once assumed to be an inevitable consequence of social categorization [12], social psychologists have long noted the distinction between explicit prejudice (negative affect towards an outgroup) and endorsement of stereotypes (cognitive representations of culturally held beliefs about a group) [1]. Whereas less research has focused on the affective–cognitive distinction in implicit cognition, implicit evaluations of social groups (implicit

*Correspondence: zoe.liberman@psych.ucsb.edu (Z. Liberman).
prejudice) may also be distinct from implicit stereotyping, and these constructs have separable influences on human social behavior [18]. Nonetheless, there are many important open questions about each of these outcomes of social categorization. For example, what is the time-course of affective and categorical thinking about groups [19], and how does one influence the other? How do social categories work similarly to, and differ from, non-social categories [11], and how do stereotypes about groups develop in the first place [20]?

One way to continue to answer important questions about the nature of social categorization is to look to developmental psychology. For example, research with infants and children can ask whether social preferences and inductive inferences based on social group membership always emerge together, or whether they arise separately [21]. Findings from such studies can demonstrate whether prejudice and stereotyping are inevitable consequences of dividing the world into conceptually-meaningful social categories, or whether humans are able to use group divisions in meaningful ways without these negative outcomes. We examine new experiments with children and infants that address the nature and origins of the human capacity to form social categories. Considering social categories from a developmental perspective does more than merely shed light on when social categorization and its downstream consequences arise -- it can also reveal the cognitive processes that shape the human ability to form social categories and provide insight into how negative consequences of social categorization begin and might be mitigated.

Social Categorization in Childhood
Social preferences for members of one’s own social group, and rich conceptual inferences based on social group membership, are in place by the time children enter formal schooling. For example, children have both explicit and implicit preferences based on gender, race, and linguistic group [22–25]. Children also look to ingroup, rather than outgroup, members when learning new information [26–30], show partiality towards the ingroup when allocating rewards and punishments [31,32], and acquire negative stereotypes associated with social group membership [20].

Recent research indicates that children use social categories to make productive social inferences. For example, children expect members of a social group to share deep properties, including preferences, traits, and norms [33–37], and they expect characteristics that mark social-category membership to endure over time [38,39]. Indeed, preschoolers expect group members to follow social conventions [40], and negatively evaluate people who do not follow the conventions and norms of their social group [41,42], suggesting that they view conforming to the group as a fundamentally important feature of group membership. In addition to sharing common attributes, members of a social category are typified by a rich relational structure, such that social categories support inferences about patterns of interpersonal interaction. For example, by early childhood, people expect members of a social category to be loyal to one another, to engage in prosocial relationships, and to share resources with each other [43–46]. In fact, whereas children think that people in a social group must refrain from harming one another, this expectation does not always hold between members of different social groups [45].

Although children hold intuitive theories that social categories are natural kinds and that social categories should mark social obligations [47], children apply these two intuitive theories differently to different social groups. Children treat gender as a natural kind [48], but they do not view novel groups [49,50] or race [51] as marking fundamental similarities between category members. In addition, children of the same age do use novel groups and race for predicting patterns of social interaction [46,51]. Children may initially view social categories as marking social obligations, and later come to see these categories as natural kinds [47].
Box 1. Infants May Prioritize Informative Social Categorization Signals

Instead of varying across communities based on learning which dimensions carry the most functional relevance (Box 3), the earliest social categories of infants appear to prioritize features that have fundamentally signaled social group membership across human evolutionary history. A prioritization account could help to explain potentially counterintuitive research on race. Because people in hunter-gatherer bands likely never traveled far enough to encounter someone of a different ‘race’, race might not be a prioritized signal of social group [100]. Indeed, although infants perceive race (Box 2), and children prefer own-race social partners [22], children do not use race as a conceptually-rich category. Children do not automatically encode race [3], do not make race-based inductive inferences [46], and do not always expect race to be stable [38,87]. Instead, seeing race as relevant for social categorization depends on social experience: minority race children, who likely think and talk more about race, see race as a defining feature of social identity earlier in development than do majority race children [38,87]. In addition, growing up in racially diverse areas decreases children’s racial essentialism [101], and racial essentialism leads children to treat ambiguous race faces as outgroup members [102], suggesting that exposure to diversity could decrease prejudice. By contrast, gender may be a dedicated dimension of social categorization: children automatically encode gender [3], and make gender-based inductive inferences [38]. In fact, transgender children express clear gender identities and use gender to carve up the social world [103], suggesting that attention to gender is present across a variety of experiences and backgrounds.

If the human system for reasoning about social categories is structured to attend to evolutionarily relevant groups, which features would infants prioritize? Spoken language and food preferences vary across groups, and are constrained by sensitive periods for learning [104–107] (Box 4), making them potential candidates for prioritized social categorization. Indeed, infants see shared language and shared foods preferences as providing information about social obligation and inductive generalization [76,77,80]. Crucially, language and food preferences may have a special significance – infants do not use highly similar cues, such as object preferences, to make the same types of social inferences [80]. This account would make the further predictions (not yet tested) that the social inferences made by children would be guided by other fundamental markers of social group membership, such as kinship or knowledge of group rituals, but not by arbitrary dimensions of similarity that did not mark social group across human evolutionary history.

Alternatively, children may prioritize the significance of some social categories versus others, and reason about prioritized categories as natural kinds and as marking patterns of social interactions at earlier ages (Box 1). In either case, the available data suggest that the formation of a social group in and of itself does not inherently lead to stereotyping or prejudice: children can know about a social division, such as race or novel group, without using group boundaries to make inductive inferences [46,50], and without expressing group bias [52].

The Relationship Between Social Preferences and Social Categorization

The growing body of research on children’s social categories brings into focus the distinction between conceptually-rich belief systems about members of particular social categories (perhaps relevant to later stereotyping) and social preferences for people who are members of those categories (perhaps relevant to later prejudice). Although both develop by early school years, and they are often coincident, these two processes are not identical: they may emerge and interact in different ways over the course of development. In fact, there are many theoretical reasons to expect that social preferences could be separable from rich knowledge of social groups.

First, preferences can exist in the absence of knowledge about groups. For example, preferential looking-time methodologies, which measure infants’ spontaneously looking to a pair of faces, find that infants spend more time looking at attractive compared to unattractive faces [53], native speakers compared to foreign language speakers [24], and own-race compared to other-race faces [54]. Although preferential looking-time studies have been taken as evidence for an early-developing own race bias, preferential looking does not necessarily indicate categorization, or a preference for the ‘ingroup’. Indeed, few people would expect that looking longer at symmetrical faces could provide evidence that infants form a conceptual category of ‘attractive people’, and expect attractive people to share common essentialized properties. Instead, infants may prefer individuals with symmetrical faces for a variety of reasons, including that symmetry may indicate health [55]. Indeed, infants could look longer at symmetrical faces without grouping these faces into a category at all. Similarly, even social preferences that seem
more plausibly relevant to early ingroup bias, such as the tendency of infants to preferentially interact with speakers of their native language [24], could emerge based on liking to approach relatively more familiar social partners, without having any abstract categorization of ‘native speaker’ or ‘foreign speaker’, or even of ‘like me’ and ‘not like me’. Some looking-time data, such as when an infant who is habituated to one type of face (e.g., gender or race) subsequently looks longer at a face of someone from a different group, may be better evidence for categorization, although this categorization could nonetheless be perceptual rather than conceptual ([21,54] and Box 2 for literature review).

Second, this differentiation could function in the reverse direction: humans may expect social group membership to influence the traits of novel individuals and their patterns of social interaction independently of forming preferences or dispreferences for groups. As example, someone could use the group identity ‘Italian’ to infer properties of a person who belongs to that social group, such as what language she might speak, what foods she might prefer to eat, what religion she might practice, and which other people she might interact with. Someone could also make these inferences about a person holding a different group identity, such as ‘Japanese’. Although humans may have a tendency to automatically prefer their own group to all other groups [14], they could nonetheless make productive inferences about people from each of these two ‘outgroups’, without necessarily preferring one outgroup to the other.

At some point in development, social preferences and social categorization appear to operate in close coordination, and it is from this coordination that negative stereotypes and other negative consequences of social categorization may be forged. One possibility is that one of these processes gives rise to the other. For example, early preferences, perhaps based on familiarity (Box 2), may set the stage for the later growth of conceptual social categories.

**Box 2. Do Visual Preferences in Infancy Reflect Social Categorization?**

Infants show clear visual preferences for people from particular social groups [21,54] for review. For example, infants prefer to look at female faces [55,109] and at own-race faces [109]. These effects are due to familiarity and vary based on contact [55,109]. For example, infants who regularly see faces of diverse races do not show own-race preferences [110], and the own-race visual preferences emerges earlier for female faces than for male faces [111], suggesting that the preference is based on liking to look at the type of face that they encounter most often in their environment (own-race females).

Further, infants are better at recognizing individual novel own-race faces compared to other-race faces, and show the ability to form perceptual categories based on race [112,113]. As with visual preferences, these benefits are likely based on expertise for processing familiar faces: exposing infants to other-race faces in the laboratory can eliminate own-race facial recognition advantages [114,115]. Thus, infants’ visual responses to social categories (in terms of preferences and perceptual categorization) reflect adaptive learning about regularities in their social environment.

Are perceptual categories linked to conceptually-rich social knowledge or social expectations about category members? For example, does better categorization of own-race faces indicate expectations that members of the own-race group will share deep properties or socially interact? To date, the closest evidence comes from a recent paper reporting that infants associate own-race faces with happier music [116]. Although this finding could be relevant to early social bias, particularly because infants see music as social [117,118], it could also be due to familiarity without any social grouping or social expectations: infants have more exposure to own-race faces and positive music than to other-race faces and negative music. Thus, more research will be necessary to ask whether perceptual categorization reflects conceptually-rich social expectations.

Even if perceptual categorization alone cannot be taken as evidence for conceptually-rich social categorization, it may scaffold the complex social reasoning of infants. For example, a tendency to pay more attention to racial ingroup members could bias children towards learning only from own-race teachers [28], and seeing ingroup members as more relevant sources of information. In addition, paying less attention to outgroup members could lead to outgroup homogeneity [119,120], whereby people might view outgroup members as more similar to one another. Future research will be necessary to explore how early differences in visual attention may relate to later-emerging conceptualizations of the social world.
Box 3. Using Functional Relevance To Form Social Groups
Because any dimension could theoretically be (or become) meaningful for social categorization in a particular community, infants and children may be ready to detect groups based on any feature, if given the appropriate input [88]. Indeed, decades of research have indicated that humans show preferential treatment of people to whom they share only a ‘minimal’, similarity. For example, people prefer ingroup members even when the group is assigned and is based on an arbitrary (and untrue) personality feature, such as being ‘overestimators’ [14]. Preferring ‘minimal’ ingroup members begins early in childhood. In both classic experiments, such as the robber’s cave [121], and more modern research [44,122–124], children prefer other people who are in their randomly assigned group.

Preferences for minimal ingroup members likely do not arise because people think that the groups are random and meaningless, but instead could emerge because participants believe that they share important features with others in their group (e.g., thinking that ‘overestimators’ are more similar to one another than they are to ‘underestimators’), or they believe that the groups must be functional because they are labeled and used by a figure in power (e.g., the experimenter). Indeed, drawing attention to the relevance of a category through labeling, generic statements, and functional use increases the likelihood that children will form preferences for their minimal ingroup [65], increases their propensity to use minimal category membership to make inferences about the behaviors of individuals [125], and heightens their expectations that members of the same minimal group will share essential similarities [66,67].

Together, these studies elegantly demonstrate that ‘minimal’ characteristics, which are not typically seen as crucial in our society, can become relevant when attention is drawn to them in a laboratory context. However, although children form social preferences for minimal ingroup members, they show stronger group-level inferences and higher levels of own-group biases when reasoning about less arbitrary categories such as gender [31]. Future work will be necessary to determine whether these differences are due to the fact that children likely have more experience of seeing these less arbitrary categories, such as gender, being used functionally in their communities, or because certain categories are more readily used for social categorization regardless of a child’s particular experiences (Box 1). One type of research that may help to resolve this debate would be work that investigates social cognition of children attending gender-neutral preschools, who likely hear less-gendered generic language and who likely see less-gendered division of labor [126].

Alternatively, children may quickly detect the category structure in the social world, and prejudice and stereotypes may result when category-based knowledge combines with children’s self-categorization and cues from society about the importance of social categories (Box 3). In contrast to each of these views, we propose that social preferences and rich category-based beliefs emerge in parallel early in development, and may not inevitably interact to form prejudice.

The Origins of Social Preferences and Social Categorization in Infancy
The bulk of research on early social reasoning by infants focuses on early emerging visual and social preferences. As discussed earlier, infants spontaneously look longer at attractive faces, female faces, own-race faces, and faces of native language speakers [24,53,54,56], and babies show a familiar-race bias in face perception (Box 2). In addition, infants are more likely to approach, interact with, and imitate people who share their preferences or who speak their native language [24,57–63]. Although these visual and social preferences could be signs of early emerging bias and prejudice, these preferences may instead operate completely differently from adult prejudice. For adults, ingroup favoritism is based on liking someone and rewarding them specifically because of their membership in the ingroup. That is, the partiality of adults towards the ingroup is depersonalized: it applies to all group members and does not depend on the perceiver being known by, or being related to, the target [64]. The social preferences of infants, by contrast, could arise solely based on an affinity for more familiar individuals. In fact, infants could prefer particular individuals over others, without grouping preferred individuals into categories at all. Thus, although social preferences may have functional value, by guiding infants towards relevant social partner [26], they do not per se indicate that infants are reasoning about people as members of conceptually-rich social categories. Other evidence will be necessary to demonstrate that infants could form inductively-useful social categories.
Indeed, it is theoretically plausible that the ability to reason in sophisticated ways about people as members of social categories arises slowly and depends on children acquiring information about members of social groups through observation of other people’s behaviors, and through older individuals’ explicit teaching and testimony. That is, children’s social categories could be grounded in the stereotypes and beliefs of adult members of their social and cultural community. Indeed, input from adults clearly influences children’s reasoning about social categories: children are more likely to see minimal social categories as informative when adults consistently label and use the categories functionally [65]. Hearing generic language about social categories leads children to be more likely to form a novel category [66], and to reason in essentialist ways about members of the novel group [67]. Such input could lead infants to move from forming preferences for familiar individuals, to forming adult-like preferences for people based on their identity as members of different social categories.

Alternatively, the ability to form relationally embedded social categories with inductive potential could plausibly be in place very early in life. A recent surge of research – outside the visual and social preferences of infants – provides evidence that infants have the cognitive capacities that may underlie conceptually-rich social categorization. Infants can think about individual items as members of conceptual categories [68], form inductive inferences [69,70], and track complex social relationships [71–79]. Below we review evidence suggesting that conceptually-rich social categorization emerges early in life.

**Looking-time Studies Provide Evidence for Early Conceptually-Rich Social Reasoning**

Research using violation of expectation looking-time methodologies, which assess the responses of infants to the actions and interactions of others, can provide a clearer view than can measures of preference into the ability of infants to form conceptually-rich social categories. In particular, because looking-time studies ask about third-party expectations, these measures can elucidate whether infants reason in abstract ways about people as members of social groups regardless of any familiarity preferences. Violation of expectation studies on infants’ understanding of social groups evaluate whether infants use cues of group membership to form expectations about the attributes and interactions of other people. For example, in one set of studies, infants inferred that characters who moved in synchrony would subsequently perform the same action [78], suggesting that they made the inductive inference that belonging to a group would influence the likely behavior of each group member.

One particularly illustrative test case of using violation of expectation studies to investigate the ability of infants to form abstract social categories comes from research on language as a marker of social group (Box 4). In these studies we showed infants two native bilingual actors who were presented as members of the same group (either two English speakers or two Spanish speakers) or as members of different groups (one English speaker and one Spanish speaker). We then asked how infants used the information about the actors’ language to inform their expectations about the interactions and attributes of the actors. In one study, we asked whether infants expected the actors to affiliate with one another or to socially disengage from one another. The responses of the infants varied based on the actors’ group membership: infants who saw both actors speak English looked longer when they disengaged, suggesting that they expected affiliation, whereas infants who saw the actors speak different languages looked longer at affiliation, suggesting that they perceived affiliation as being unexpected [76]. Thus, like adults and older children [45], infants expect people who speak a common language to engage, but they do not hold these same social expectations for people who speak different languages. These data suggest that infants view social relationships as being embedded in broader shared social categories.
Box 4. Language Is a Potent Cue to Social Structure

Research on intergroup cognition often focuses on race, gender, and age. Indeed, adults and children are sensitive to these social categories and use membership of them to guide their preferences and learning [2]. Nevertheless, despite a wealth of evidence from the neighboring social sciences disciplines of linguistics and anthropology that language and accent serve as particularly reliable signals of social group membership [104,105], and that attention to language can surpass attention to visual cues in social categorization tasks [127–129], language is often overlooked in social psychology research on intergroup cognition [130].

The sociality of language emerges early: infants prefer native-language speakers [24], and infants and children look to linguistic ingroup members to learn new information [26,29,57,60,62]. In fact, infants use language for more than personal decisions about whom to like or whom to learn from; they create conceptually-rich social categories based on language, whereby they use language to make predictions about people’s likely traits and social interactions. For example, they expect same-language speakers to be more likely to affiliate [77], and expect same-language speakers to share important social similarities, even when the language shared by the speakers is unfamiliar to them [80].

Language may inherently mark social group, and using language to divide the world into groups continues across development. Indeed, by preschool, children expect native speakers, but not foreign speakers, to follow social conventions [40], and children acquire linguistic stereotypes [131,132]. Does early language-based categorization feed into xenophobia? If so, which experiences mitigate these biases? In our research, infants raised in bilingual environments generalized information even across different-language speakers [80], suggesting that multilingual exposure may cause people to be less likely to see language as marking concrete social boundaries. More work will be necessary to determine whether experiencing a diverse linguistic community could reduce stereotyping and prejudice, and, if so, whether the reduction in bias would be specifically in the domain of language (e.g., by increasing attitudes towards foreign speakers) or would be broader. These questions about the role of experience in bias formation and reduction should also be asked about other social categories, including those typically studied (gender, race, and age) and those that are less studied but of potentially high evolutionary relevance (Box 1).

In another series of studies we asked about the inductive generalizations of infants aged 11 months. After being presented with same-language or different-language speakers, infants were shown one speaker’s food preference. Subsequently, infants were shown the second speaker disagreeing with the first speaker (by disliking the previously liked food), or expressing a negative opinion of a previously uneaten food. Infants selectively generalized information across same-language speakers: they looked longer at the disagreement when the actors spoke the same language but not when the actors spoke different languages, suggesting that they found it unexpected for people from the same social group, but not all people, to disagree [80] (Figure 1). Infants show a similar pattern of inductive generalization of labels: they expect people who speak the same language to use the same novel labels to refer to the same object [81], but do not expect people who speak different languages to use the same object labels [82].

Interestingly, infants’ inductive inferences based on shared language are not limited to speakers of a familiar language: infants are equally likely to generalize information across same-language speakers when the shared language is the infants’ native language (English) and when the shared language is unfamiliar to the infant (Spanish) [80]. Thus, infants’ expectations did not require any specific information or experience with that linguistic group to infer that people who speak the same language share relevant similarities.

An Initial System for Social Categorization in Infancy

Taken together, the findings from violation of expectation studies suggest that infants can generalize information selectively across same-language speakers, and make inferences about social relationships based on language. Therefore, at least in the case of language, social categorization by infants shares crucially important features of older children’s and adult’s social categorization: they use information about group membership to infer whether people will share properties, and how people will interact. Thus, conceptually-rich social categories emerge before verbally provided information can affect social knowledge, suggesting that the ability to form social categories does not depend on explicit learning about the cultural or stereotypic content associated with different groups, and that the ability to use these categories
to draw inferences about social structure likely drives social thinking and learning from early in ontogeny.

Although this recent work on inductive generalization by infants and their inferences about social relationships focuses on language categories, we hypothesize that infants could apply these same abstract features of social categorization to other groups that they think are socially important. Specifically, infants may have a system for thinking about people as members of social groups that is present early in ontogeny, such that infants are ready, early on, to apprehend and generalize information across individuals in a social category. Of course, children’s social category knowledge grows significantly across development, and social partners play an important role in this process. In particular, we expect that social experiences (such as the typical environment of infants and the information they receive from social partners) would modulate which features infants would see as relevant markers of social categories.

We hypothesize that infants begin by seeing specific features of human behavior as fundamentally relevant to social categorization (Box 1), and, based on their social experiences, they learn to update the set of features they use to divide the social world into groups. Under this account, infants could require different experiences to form different social categories. That is, infants may initially expect that shared features that have defined group membership across human evolutionary history, such as language (Box 4), food preferences [80,83,84], and

---

Figure 1. Infants’ Use of Language when Making Inductive Inferences. This figure details methods and results from [65]. Monolingual infants generalized food preferences across same-language speakers, finding it unexpected when they disagree, but did not generalize food preferences across different-language speakers. By contrast, infants from bilingual backgrounds generalized even across speakers of different languages.
engagement in ritualistic actions [85], would mark people as members of a social category. By contrast, infants would likely not see an arbitrary similarity, such as being randomly assigned to wear the same color mittens [61,86], as defining membership in a conceptually-rich social group.

However, with experience, infants and children likely update the list of dimensions that are seen as relevant for social categorization. Thus, although humans might have predispositions to attend to some markers of social division over others, the features that are relevant in each infant’s and each child’s community will certainly impact on social categorization across development. For example, experience with group norms can lead children to form social categories based on dimensions that were not relevant in our evolutionary past, such as race (Box 1). Demonstrating the importance of social experiences on social thinking, minority race children, who likely have more experience of thinking about race, reason about race as an important social category at earlier ages than do majority race children [38,87]. These ideas are consistent with developmental intergroup theory (DIT), which suggests that any dimension that is marked and made salient in a child’s community (e.g., by explicit input from important social partners) may be able to be co-opted into the human system for reasoning about abstract social categories [88]. Indeed, this process likely underlies minimal group effects: researchers approximate social relevance by highlighting an arbitrary similarity, leading people to use the arbitrary feature as they would use an important social group marker (Box 3).

Malleability in the features that are seen as relevant for social categorization may also work in the reverse direction – such that even categories that served as fundamental social group markers may be abandoned based on early social experiences. As an example, although we argue that language reliably marks social group, and may be prioritized in early social categorization by infants, differences in sociolinguistic environments of infants may influence whether they reason about language as marking social categories. Whereas infants from monolingual environments refrain from generalizing information across people who speak different languages, suggesting that they may view different-language speakers as members of distinct social groups, infants from multilingual backgrounds generalize even across different-language speakers [80]. Therefore, infants who regularly see people who speak diverse languages interact may be less likely to use spoken language as a boundary for social groups. Thus, variations in important features of social environments could impact on broader reasoning about the social world. Future research should ask how social experiences influence categorization on both potentially prioritized dimensions and on dimensions that humans may learn are important via social transmission (see Outstanding Questions).

Using Malleability of Social Categorization To Reduce Social Prejudice

Although human infants may be ready to form conceptually-rich social categories, the fact that forming generative inferences based on category membership can, in theory, be separated from dislike of the outgroup [1], and the fact that the particular dimensions upon which humans form social categories are malleable [38,80], suggest that prejudice against members of particular social groups is not inevitable. Developmental research sheds light on the relationship between categorization and preference formation, suggesting that studying human reasoning across the lifespan is crucial for understanding the emergence and malleability of intergroup bias. Specifically, important future studies will continue to investigate how and when social preferences and conceptual reasoning about social groups come to operate together, leading to prejudice and discrimination. One possibility is that children first form simple preferences for familiar people, and then later generalize these positive associations to personally unfamiliar members of their broader social group, leading them to show ingroup positivity [64] and eventually outgroup negativity [52]. Alternatively, children may form conceptually-rich social categories, and then come to self-identify with one category, at which point they may begin to
show adult-like depersonalized preferences for members of their own group, leading to bias (see Outstanding Questions).

One particularly important area for future study involves investigating how parents and educators can limit the transmission of bias. One prominent way that adults transmit information about social categories is through their language. For instance, generic language refers to groups rather than individuals (e.g., ‘boys like X’, or ‘Hispanics live in Y’), signifying that groups are enduring, highlighting group differences, and teaching children that the group distinction is meaningful. Indeed, hearing generic statements about a novel social group increases the likelihood that children form a conceptually-rich social category, and can lead children to develop essentialist thoughts and stereotypes about the novel group [65–67]. Thus, parents and educators may strive to speak about people as individuals (e.g., ‘This boy likes X’), instead of speaking about whole categories of people, so as to reduce essentialist tendencies.

As a caveat, it is impossible, and potentially counterproductive, to avoid all conversation that remarks on social group membership. Indeed, research on ‘colorblind’ interventions shows that purposefully refraining from all discussion of a category (in this case, race) can be ineffective and can even lead to increased prejudice [89,90]. Nevertheless, focusing on people as being distinctive individuals, as opposed to members of groups with collective properties, is one area in which language can be used in smart ways to potentially reduce the tendency of children to form a new conceptually-rich social category, and to lower the transmission of bias towards members of the highlighted social group. In support of this idea, introducing people to counterstereotypic individuals from a particular social group has been one of the most effective ways to reduce implicit bias for both adults [91] and children [92].

Instead of trying to halt the formation of social categorization in the first place, many interventions have focused on reducing the social significance that people ascribe to the categories to which other people belong. For example, interventions aimed at reducing prejudice based on gender and race have successfully led to reduced explicit and implicit bias, to smoother cross-group interactions, and even to increased overt actions aimed at promoting equality [93–95]. These interventions probably do not change the likelihood that adults categorize people into social groups, but instead may help to participants change the perceptions, beliefs, and stereotypes they ascribe to those social categories. Interesting future questions concern how to leverage the insight gained by studies of bias reduction among adults to create manipulations that are effective with children. Indeed, current research suggests that implicit associations are malleable based on new information [96,97], and that similar interventions may be effective for both adults and children [91,92]. Indeed, efforts to change the structure of social categorization among children may be even more impactful because children have less experience, meaning that their stereotypes and bias may be less entrenched and easier to overcome.

Concluding Remarks
Social categorization has vast implications for myriad aspects of human social life. We have presented evidence that developmental psychology can inform our understanding of the mechanisms underlying social categorization and its downstream negative consequences. To this end we aimed at providing a review of developmental research on social categorization, and we have argued that conceptually-rich social categorization is functionally different from social preferences for individual members of social groups. Separating these two processes can lead to a better understanding of the mechanisms and implications of each type of data. Although these constructs may act in tandem in adulthood, it is theoretically and empirically possible to separate them: having a social preference for people from a familiar background does not require reasoning about abstract similarities between group members, and the initial formation of social predictions based

### Outstanding Questions

**Role of self-categorization: how early do infants and children self-categorize into social groups?** Can conceptually-rich social categories exist before the development of a sense of self? How does acquiring a sense of self and self-categorization influence social categorization, social preferences, and prejudice? How does the social status of the groups to which the child self-categorizes impact children’s social learning, own-group preferences, and prejudice? What happens when children identify with more than one group?

**Role of social experience: what types of experiences influence infants’ social categorization?** Is the link between experience and social categorization specific (e.g., language experience impacts on thinking about language as a social marker) or broad (e.g., language experience leads to more-flexible categorization generally)? Are early social experiences more important than later ones? How does experience impact the tendency to create a category at all as well as the tendency to use that category to make social inferences?

**Malleability of prejudice: what types of interventions most successfully reduce prejudice?** Do the same interventions work across the lifespan? Does reducing a social preference based on one social category (e.g., race) reduce social preferences or prejudice more broadly (e.g., change gender stereotypes)?

**Priorities in social categorization: which social boundaries are infants most likely to attend to?** Are infants’ earliest social categories based on dimensions that have fundamentally marked social groups across human evolutionary history? Is reasoning about potentially prioritized social categories less malleable than reasoning about social categories that are acquired later? Are categories that are learned to be important based on social input (e.g., race) used identically to potentially prioritized categories once they are acquired?
on social categories does not obligate a preference or dispreference for a particular group. In fact, although hearing generic language increases essentialist reasoning about novel social groups \([65–67]\), children who hear generic language about a novel social group do not initially show a lower level of liking for members that group compared to children who heard specific language about members of the novel group \([98]\). Even for adults, higher levels of essentialist reasoning are not always related to higher levels of prejudice towards social groups \([99]\). Therefore, even if forming social categories and making social inferences based on these categories is a basic part of human cognition, prejudice is not inevitable.

Many open questions remain regarding the origins of social categorization (see Outstanding Questions), including questions about which dimensions infants’ see as being fundamentally relevant to social categorization (Box 1), and about how experience across the lifespan shapes the use of these social categories in real interactive contexts (Box 3). However, this growing body of research suggests that an ability to see people as members of social groups, and to use these groups to inform inferences about the social world, emerges in infancy. Clarity in the evidence deemed necessary to demonstrate conceptually-rich social categorization in infancy will propel these important inquiries forward.

Acknowledgment

We acknowledge the support of NICHD grant R01 HD070890 to K.D.K.

References
